The Goldbach conjecture: Why is it difficult?

Paul Turner
Canberra, ACT
pturner282@gmail.com

The opinion of the mathematician Christian Goldbach, stated in correspondence with Euler in 1742, that every even number greater than 2 can be expressed as the sum of two primes, seems to be true in the sense that no one has ever found a counterexample (Boyer & Merzbach, 1989, p. 509). Yet, it has resisted all attempts to establish it as a theorem.

The following discussion is intended to explain, partially and from a naïve point of view, why the search for a proof has been difficult. This is not to say a student, captivated by the prospect of an open question, should not make the attempt. The value in considering arguments that do not lead to a proof may lie in the potential for hinting at hidden depths to be explored and new approaches to be looked for.

If further justification were needed for including such a discussion in a journal mainly read by senior secondary mathematics teachers, one could point to the stated aims of the Specialist Mathematics course in the Australian Curriculum. While number theory as a topic is not included in the Australian Curriculum, the aim to develop students' "reasoning in mathematical and statistical contexts " and "ability to construct proofs" is clearly stated (ACARA, 2016). This topic in number theory certainly assists in those aims both because of its particular technical nature and also because of its capacity to capture the imagination and so provide a motivation for learning in mathematics generally.

Primes

Usually, there are several ways to decompose an even number into a pair of odd primes. One might wonder how it could be possible for an even number to exist that does not have such a Goldbach partition.

Indeed, if we restrict our attention to the first n prime numbers, we can count the number of expressions of the form $p_i + p_j$, i, $j \leq n$, with $i \leq j$ that can be formed from them, namely $\frac{n(n+1)}{2}$. The count $\frac{n(n+1)}{2}$ of prime sums up

to the nth prime includes the odd sums, of which there must be $n - 1$. After excluding these, the count of even sums must be $\frac{n(n-1)}{2} + 1$.

An example should clarify how this count comes about and will reveal a consequence. There are ten expressions $p_i + p_j$ for the primes up to $p_4 = 7$, generating the following sums:

2 + 2 = 4	2 + 3 = 5	2 + 5 = 7	2 + 7 = 9
	3 + 3 = 6	3 + 5 = 8	3 + 7 = 10
		5 + 5 = 10	5 + 7 = 12
			7 + 7 = 14

All the even numbers from 4 to $2p_4 = 14$ are in the list. In this case, not all of the sums are distinct ($3 + 7 = 5 + 5 = 10$). This is bound to occur because there are only six even numbers available, yet there are seven even sums. So, the sums cannot be distinct and there must be at least one even number between 4 and 14 with more than one Goldbach partition.

Again, up to the tenth prime number there are $\frac{10\times 9}{2} + 1 = 46$ even sums, shown in the following table:

+	2	3	5	7	11	13	17	19	23	29
2	4	5	7	9	13	15	19	21	25	31
3		6	8	10	14	16	20	22	26	32
5			10	12	16	18	22	24	28	34
7				14	18	20	24	26	30	36
11					22	24	28	30	34	40
13						26	30	32	36	42
17							34	36	40	46
19								38	42	48
23									46	52
29										58

In this case there are 28 even numbers in the set $\{4, 6, \ldots 2p_{10}\}$ since the tenth prime number is 29 and we leave out the number 2. But not all of the 28 are in the list. As before, the sums cannot be distinct because there are 46 sums but only 28 even numbers.

Note that in the example, the numbers 50, 54 and 56 are missing from the list. To express these as Goldbach sums the extra primes 31 and 37 are needed.

There appears always to be an excess of even prime sum expressions that can be formed by primes up to the nth over the number of even numbers (except 2) up to $2n$. That is, $\frac{n(n-1)}{2} + 1 > p_n - 1$.

To confirm this, at least for large enough n, the prime number theorem, conjectured by Gauss and proved much later by Hadamard and Vallée-Poussin, can be brought into play. The prime number theorem gives an approximation of the number of primes $\pi(N)$ less than a number N. It says,

$$\pi(N) \approx \frac{N}{\ln N}$$

or, in the present case, $$n \approx \frac{P_n}{\ln P_n}$$

So, we check the truth of

$$\frac{\left(\frac{x}{\ln x}\right)\left(\frac{x}{\ln x}-1\right)}{2}+1 > x-1$$

This is done below with the help of graphing software, showing the inequality to be true when $x > 17$, as shown in Figure 1.

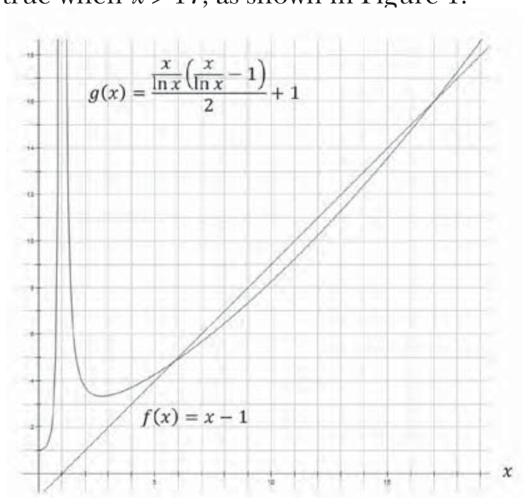

Figure 1

In fact, experiment shows the inequality $\frac{n(n-1)}{2} + 1 > p_n - 1$ to be true when $n > 3$. Thus, for $n > 3$, there must be at least one even number in the range up to $2p_n$ with more than one Goldbach partition. However, nothing stands in the way of the existence of even numbers less than $2p_n$ that are *not* in the list of sums (as was seen in the case of p_{10}).

Some questions arise for which no answers present themselves:

1. If we consider the Goldbach sums that can be formed with the odd primes up to and including p_n, what even numbers less than $2p_n$ could possibly be left out of the list of sums?
2. What further primes, necessarily between p_n and $2p_n$, are available that could be effective in forming the sums needed to complete the list?

The abundance of the prime sums $p_i + p_j$, $i, j \leq n$ makes the existence of a non-Goldbach even number seem unlikely. Yet, it does not preclude the possibility. There could be, for some n, such a large number of expressions $p_i + p_j$ all of which evaluate to the same sum, that an even number somewhere is left without a Goldbach partition. Another approach is needed.

Even numbers

We try looking at various sets of partitions of even numbers, looking for ever smaller sets that still contain the set of prime partitions.

An even number N can be decomposed into a sum of integers x and y in $\frac{N}{2}$ distinct ways; that is, the equation $x + y = N$ has $\frac{N}{2}$ distinct solutions. The Goldbach partitions of N, if there are any, are certainly contained in this initial set of partitions.

Now the only even prime, 2, belongs to a Goldbach partition in the case $N = 4$ but otherwise it cannot be part of a Goldbach partition since it would have to be paired with another even prime number. So, apart from this solitary case, Goldbach partitions must comprise odd numbers only. Thus, if x and y are restricted to the odd integers, the equation $x + y = N$ now has about half as many solutions as previously. Again, the Goldbach partitions of N, if there are any, are contained within this smaller set.

There exists a still smaller and more interesting subset of sums $x + y$ that is guaranteed to contain all of the Goldbach partitions in most cases, the exception being the special case of even numbers that are twice a prime. We construct this set by requiring that both x and y be totients of N. That is, they are numbers that do not exceed N and have no divisors in common with N other than 1. The set of totients of a number N is non-empty since it includes 1 and, for numbers greater than 2, it also includes $N - 1$.

If t is a totient of N, then by definition no prime factor of t divides N. It follows that no prime factor of the complement $N - t = t'$ divides N. Thus, the totients of N come in pairs (t, t') such that $t + t' = N$. We see also that t and t' themselves have no common factors since, if they did, the fact that $t + t' = N$ would imply that a factor common to t and t' would also divide N.

Since t and t' have no common factors, we have $t \neq t'$ and it follows that the number of totients of N is always even (whatever the parity of N). The totients thus can be formed into a set of sums $t + t' = N$ and these sums must include most of the Goldbach partitions of N. All of them are included when $N \neq 2p$ for any prime p. But, when $N = 2p$ for a prime p, we should augment the set of partitions involving totients by including the sum $p + p$.

We could perhaps then remove the totient pair $(1, N - 1)$ as it does not form a Goldbach sum, but this is unimportant.

Thus, for example, the number 14 has the following set of totient sums augmented by $7 + 7$ and reduced by the removal of $1 + 13$. It includes the two possible Goldbach partitions.

$$\{3 + 11, 5 + 9, 7 + 7\}$$

The foregoing amounts to a proof of a related theorem that falls well short of the Goldbach conjecture:

> Every even number greater than 2 can be expressed as the sum of two *relatively* prime numbers.

By an application of the inclusion-exclusion principle, it can be shown that the number of totients of a number N is given by Euler's phi-function:

$$\varphi(N) = N\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\ldots$$

where the p_i are the distinct prime factors of N. Hence, the number of totient sums, $t + t' = N$, must be $\frac{1}{2}\varphi(N)$.

Thus, we have an upper bound on the number of possible Goldbach sums for a given N. But, the Goldbach conjecture amounts to the claim that there is a non-zero lower bound. This line of enquiry has not led to the establishment of such a lower bound but it seems worth pursuing because it may help to explain some observed facts about the occurrence of Goldberg partitions.

It has been noted (Australian Association of Mathematics Teachers [AAMT], 2016) that divisibility by 6 tends to confer on even numbers a relative richness in Goldbach partitions compared with numbers not divisible by 6. For example, 2016 has 73 Goldbach partitions while only 27 prime pairs sum to 2012 and 35 sum to 2014.

This is explained probabilistically by noting that prime numbers are congruent to either 1 or –1 (mod 6). Thus, the possible sums of prime numbers written modulo 6 are

$$1 + 1\ ,\ 1 + (-1),\ (-1) + 1,\ (-1) + (-1)$$

giving results that are respectively 2, 0, 0, –2 (mod 6). On average, there are twice as many ways to form, as a sum of primes, a number divisible by 6 as there are ways to form either of the other types of even number. Therefore, we expect, although we cannot insist on it, that numbers divisible by 6 will have about twice as many Goldbach partitions as the other types of even numbers.

Unfortunately, there are exceptions. Armatys (via AAMT, 2016) cites the numbers 1540, 20 020 and 80 080 as exceptional cases, tentatively calling them *Anya numbers*. These are not divisible by 6 and yet they have similar numbers of Goldbach partitions to their near neighbours that are multiples of 6. In particular, 1540 has the same number of Goldbach partitions as 1542 and the other two have more of them than their near neighbours, 20 022 and 80 082, both of which are multiples of 6.

An explanation is called for that accounts for both the observed mod 6 phenomenon and the exceptions. The machinery introduced above, the Euler phi-function and the prime number theorem, promises to make this possible.

Recall that the prime factors of a number N are excluded from its list of totients but all the other prime numbers less than N are included.

Consider a range of even numbers surrounding some N. For example, {N – 4, N – 2, N, N + 2, N + 4}. For large enough values of N, the number $\pi(N)$ of primes less than N does not vary by much over the range of numbers in the neighbourhood of N. However, the number of totients belonging to these numbers varies considerably.

The phi-function

$$\varphi(N) = N\left(1-\frac{1}{p_1}\right)\left(1-\frac{1}{p_2}\right)\ldots$$

has a factor

$$\left(1-\frac{1}{p_1}\right)$$

for every distinct prime factor of N. Each successive factor

$$\left(1-\frac{1}{p_1}\right)$$

included in the function has the effect of reducing the size of $\varphi(N)$. We note that the amount of this reduction is greater for smaller prime factors p_i.

So, numbers have fewer totients when they have small prime factors. As a consequence, the density of prime numbers among the totients is greater when N has small prime factors. In particular, the density of primes is high when 3 is a prime factor, that is, for even numbers divisible by 6.

The higher density of primes implies an increased number of opportunities for the formation of Goldbach partitions of N. As in the earlier probabilistic explanation, a greater number of Goldbach partitions is expected but still not guaranteed. We might argue, for example, that if two-thirds of the totients of N were primes, then the probability that both components of a totient sum $t + t'$ are prime is about $\frac{4}{9}$.

If we consider the densities of primes in the totient sets over the previously mentioned range of numbers $\{N-4, N-2, N, N+2, N+4\}$, we see that every third number in such a set is divisible by 6 and therefore has a relatively high density of primes.

It is possible, however, for another number in the set to be similarly dense in primes if it has 5 and 7 among its prime factors but not 3. This is because $\left(1-\frac{1}{3}\right)$ in the phi-function is not very different from $\left(1-\frac{1}{5}\right)\left(1-\frac{1}{7}\right)$. This accounts for the Anya numbers.

By way of illustration, consider the even numbers 70, 72 and 74.

$70 = 2 \times 5 \times 7$. From these 3 prime factors we find

$$\varphi(70) = N\left(1-\frac{1}{2}\right)\left(1-\frac{1}{5}\right)\left(1-\frac{1}{7}\right) = 70 \times \frac{1}{2} \times \frac{4}{5} \times \frac{6}{7} = 24 \text{ totients}$$

and these include 16 odd primes. Hence, the density of primes among the totients is about 0.66.

We see that there are 12 totient sums $t + t' = 70$. Thus, at least $16 - 12 = 4$ of them have to be Goldbach partitions. In fact, there are five. They are: $\{3 + 67, 11 + 59, 17 + 53, 23 + 47, 29 + 41\}$.

Probabilistically, we expect $0.66^2 \times 12 \approx 5.2$ Goldbach sums.

$72 = 2^3 \times 3^2$ has

$$\varphi(72) = 72 \times \frac{1}{2} \times \frac{2}{3} = 26 \text{ totients}$$

including 17 odd primes. The density of primes among the totients is approximately 0.65.

There are 13 totient sums $t + t' = 72$. Therefore, at least $17 - 13 = 4$ of them must be Goldbach partitions. In fact there are six, namely:

{5 + 71, 11 + 61, 13 + 59, 19 + 53, 29 + 43, 31 + 41}.

Probabilistically, $0.65^2 \times 13 \approx 5.5$ Goldbach sums are expected.

$74 = 2 \times 37$ has

$$\varphi(74) = 74 \times \frac{1}{2} \times \frac{36}{37} = 36 \text{ totients}$$

including 18 odd primes. The density of primes among the totients is 0.5.

There are 18 totient sums $t + t' = 74$. Therefore, the 18 primes could conceivably be distributed in such a way that there is no totient sum with both components prime. Nevertheless, there exist four Goldbach partitions in this case: {3 + 71, 7 + 67, 13 + 61, 31 + 43}.

Probabilistically, we expect $0.5^2 \times 18 = 4.5$ Goldbach partitions.

In the illustration, a pigeonhole principle argument has been used tacitly to show that when the primes less than an even number are sufficiently dense among its totients, the existence of Goldbach partitions is unavoidable.

More precisely, the existence of Goldbach partitions is assured when the density of primes among the totients of an even number is greater than 0.5. At lower prime densities, the probabilistic argument leads to the expectation of a non-zero number of Goldbach partitions but it does not make such an occurrence necessary.

Where not to look for a counterexample

The obvious case $N = 2p$ has already been mentioned. Such numbers always have the Goldbach partition $p + p$. Likewise, even numbers with a higher than 0.5 density of primes among the totients are guaranteed to have a Goldbach partition.

If there is a counterexample, it is in the region where the prime/totient density is below 0.5. To show where numbers of this sort are to be found, we might construct sequences of even numbers by changing the prime factors in a systematic way using schemes like the following.

Suppose a sequence begins with a number that has a certain collection of prime factors. We might form a geometric sequence by increasing the multiplicity of just one of the factors by one at each step. The number of totients of each successive term also increases geometrically. However, the

number of primes less than each term increases at a slower rate because large primes are distributed more sparsely than smaller ones. Thus, the density of primes in the totients decreases and can be made smaller than any chosen small number by going far enough along the sequence.

For example, consider the sequence {6, 12, 24, 48…}. There are two distinct prime factors in each term and in each successive term the multiplicity of the factor 2 increases by one. The values of the phi-function for the numbers in this sequence go according to the sequence {2, 4, 8, 16…}, doubling at each step. The number of available primes at each step increases at a slower rate: {1, 3, 7, 13, 22, 42, 76, 135, 242…} so that eventually the density must fall below 0.5. This occurs after the number $768 = 2^8 \times 3$. While 768 must have at least 7 Goldbach partitions, nothing so far guarantees that the next number, 1536, will have any at all.

Again, suppose a sequence begins with a term that has a certain collection of prime factors. We might construct the sequence by introducing an additional distinct prime factor to each successive term (each larger than the one added to the previous term). The phi-function values for this sequence are easily computed when it is realised that the function is multiplicative for arguments that have no common factors. That is, $\varphi(pq) = \varphi(p)\varphi(q)$ for p, q relatively prime. In this situation, as before, the number of primes increases at a slower rate at each step than the number of totients, and the density falls inexorably.

In the following example, the sequence has the next prime number added to its list of distinct prime factors at each step: {2, 6, 30, 210, 2310, 30 030, 510 510…}. Thus the phi-function values are: {1, 2, 8, 48, 480, 5760, 92 160…}.

Again, the number of available primes {0, 1, 7, 42, 338, 2912, 8062…} increases at a slower rate at each step than the number of totients. The density drops below 0.5 after $30\,030 = 2 \times 3 \times 5 \times 7 \times 11 \times 13$ and nothing conclusive can be said, using the methods described, about the next term, 510 510.

Intuitively, we might cling to the hope of finding a counterexample among immensely large numbers where the prime/totient density is low. But, while the probability that a particular totient sum will be a Goldbach partition might become very low, the increasing number of totients belonging to increasingly large even numbers ensures that the absolute expected number of Goldbach sums also grows without bound.

For example, we find with the help of the online software resource Wolfram Alpha, that the number

$261\,782\,208 = 2^2 \times 13 \times 281 \times 26\,591$

has 120 852 480 totients and 14 339 661 primes less than it. The density of primes in the totients is therefore 0.11865 leading to an expectation of

$$0.11865^2 \times \frac{120\,852\,480}{2} \approx 850\,731$$

prime partitions. At least one of these expected partitions exists, namely: 261 782 167 + 41.

Progress

Goldbach enthusiasts search for counterexamples among very large numbers using powerful computing tools. In a way, it would be disappointing if a counterexample were to be found by such a method. More satisfactory would be a direct proof of the original conjecture. The methods described in this discussion are clearly too blunt for the problem. They seem powerless to decide the issue one way or the other. If there is ever to be a resolution, or even a partial discovery about a simpler related problem, it will be with the aid of properties and insights not broached here.

References

Australian Association of Mathematics Teachers [AAMT]. (January 2016). 2016 series. *AAMT discussion list* [email communications by J. Ousby, R. Cross, S. Thornton, A. Edwards & S. Armatys].

ACARA (n.d.). *The Australian Curriculum: Specialist mathematics*. Retrieved 2 October 2016 from http://www.australiancurriculum.edu.au/seniorsecondary/mathematics/specialist-mathematics/rationaleaims.

Boyer, C. B. & Merzbach, U. C. (1989). *A history of mathematics* (2nd ed.). New York, NY: Wiley.

Wolfram Alfa: Computational Knowledge Engine (n.d.). Retrieved 29 January 2016 from https://www.wolframalpha.com